

ESCOLA SUPERIOR ABERTA DO BRASIL - ESAB
CURSO DE PÓS-GRADUAÇÃO LATO SENSU EM
SISTEMAS DE TELECOMUNICAÇÕES

BRUNO MARIANI DE MELO

ÁUDIO SOBRE IP

VILA VELHA - ES
2011

BRUNO MARIANI DE MELO

ÁUDIO SOBRE IP

Monografia apresentada ao Curso de Pós-Graduação em Sistemas de Telecomunicações da Escola Superior Aberta do Brasil, como pré-requisito para obtenção do título de especialista em Sistemas de Telecomunicações, sob orientação da Professora Luciana Genelhú Zonta.

VILA VELHA - ES

2011

BRUNO MARIANI DE MELO

ÁUDIO SOBRE IP

Monografia aprovada em de de

Banca Examinadora

VILA VELHA - ES

2011

RESUMO

Palavras-chave: Protocolo IP, Áudio, Multicast.

Este trabalho tem como objetivo a pesquisa exploratória sobre as características, arquitetura, funcionamento e aplicações dos sistemas de áudio sobre redes IP, que surgem como uma nova alternativa para a distribuição de áudio e se apresentam como a nova geração dos sistemas de sonorização. Trata das principais tecnologias e equipamentos envolvidos e os requisitos necessários para a implementação de um sistema de áudio sobre IP, bem como as vantagens e benefícios que esse tipo de sistema pode proporcionar em relação à transmissão e distribuição de áudio no tradicional modelo analógico. Inicialmente, são apresentados os conceitos básicos de áudio e redes IP, cujos fundamentos são essenciais ao bom entendimento do texto principal que se segue, onde são explorados os aspectos técnicos e estruturais dos sistemas de áudio sobre IP. Nessa etapa, é feita uma análise dos principais componentes de um sistema típico, os arranjos e topologias mais utilizados, os métodos de codificação, protocolos, comparativo de custos, normatização internacional, entre outros aspectos.

SUMÁRIO

1	Introdução	7
2	Conceitos de Áudio	8
2.1	Áudio Analógico	8
2.1.1	Frequência.....	9
2.1.2	Amplitude.....	10
2.1.3	Decibel (dB).....	11
2.2	Áudio Digital	13
2.2.1	Digitalização do Áudio	13
2.2.2	Taxa de Amostragem	15
2.2.3	Resolução.....	15
2.3	Sistemas de sonorização	17
3	Redes Ethernet/IP	21
3.1	Modelo OSI	22
3.2	Protocolo IP.....	23
4	Áudio sobre IP.....	25
4.1	Componentes do Sistema	26
4.2	Argumentos para Áudio sobre IP	29
4.3	Infraestrutura de Rede	31
4.4	Streaming x IP.....	32
4.5	Voz sobre IP x Áudio sobre IP	33
4.6	Instalação e Custos.....	34
4.7	Topologias e Aplicações	37
4.7.1	Snake	37
4.7.2	Studio-to-Transmitter Link (STL).....	38
4.7.3	Estrela - Pequena Escala	38
4.7.4	Daisy-Chained	39
4.7.5	Estrela – Larga Escala.....	40
4.8	Padronização EBU N/ACIP	41
4.8.1	Protocolos de Transporte	41
4.8.2	Codificação de Áudio.....	44
4.8.3	Tipos de Conexões IP	46

4.8.4	Protocolos de Sinalização	48
4.9	Qualidade do Áudio.....	49
5	Conclusões	51
6	Referências	52

1 INTRODUÇÃO

Com o advento das tecnologias digitais e o desenvolvimento das redes de telecomunicações, os tradicionais sistemas de sonorização e transmissão de áudio estão passando por um profundo processo de mudança. Trata-se da migração dos antigos sistemas analógicos para os modernos sistemas digitais de áudio baseados em redes de comunicação IP, que proporcionam maiores flexibilidade, controle, qualidade, segurança e redução de custos. Essa transição tem provocado uma revolução na maneira como o áudio é gerenciado e distribuído, trazendo enormes benefícios aos que trabalham com esse tipo de mídia, como estúdios de gravação, emissoras de rádio e televisão, empresas de telefonia, instituições governamentais, entre outros.

Este trabalho aborda a tecnologia de áudio sobre IP e suas características e aplicações nos sistemas de sonorização e transmissão de áudio. Propõe-se a examinar o processo de substituição dos tradicionais sistemas analógicos pelo novo modelo baseado em redes de comunicação entre dispositivos através do protocolo IP. Nesse sentido, tem por objetivo: apresentar informações teóricas sobre conceitos básicos e introdutórios de áudio e redes IP, essenciais para o bom entendimento do conteúdo abordado; identificar as características, requisitos, vantagens e desvantagens da utilização dessas redes nas aplicações de áudio profissional; descrever os diferentes tipos de arquiteturas, topologias, métodos de codificação, compressão e transmissão de áudio; e apresentar exemplos típicos de aplicação desses sistemas nas áreas de sonorização e distribuição de conteúdo sonoro.

A metodologia adotada neste estudo é do tipo exploratória, por meio de pesquisas e levantamento de informações técnicas em livros, artigos, manuais e sites dos fabricantes de equipamentos.

2 CONCEITOS DE ÁUDIO

2.1 ÁUDIO ANALÓGICO

Entende-se por *Áudio* “[...] qualquer fenômeno no qual ocorram de 20 a 20 mil ondas por segundo” (Valle, 2007, p.9). Um sinal de áudio, em outra definição, é um “[...] sinal elétrico (tensão e corrente) que carrega informação sonora convertida por um transdutor” (Ratton, 2004, p.137). Baseado nessas definições, pode-se dizer que um sinal de áudio analógico é a representação de uma ou mais ondas sonoras convertidas em um sinal elétrico contínuo de tensão e frequência variáveis. Dessa forma, as variações de amplitude e frequência do sinal elétrico correspondem analogamente às variações da pressão do ar produzidas pelo som, como ilustra a Figura 1 abaixo. O sinal de áudio é, portanto, a essência de qualquer sistema de sonorização, incluindo os de áudio sobre IP, pois carrega a informação principal que se deseja transmitir, reproduzir e/ou armazenar.

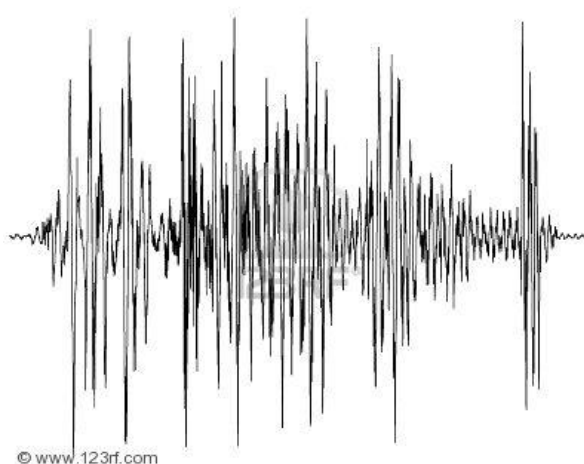


Figura 1: Representação de um sinal de áudio analógico.

Fonte: <http://www.123rf.com> (2010).

O microfone é um exemplo de dispositivo transdutor que converte as ondas acústicas em sinais elétricos, enquanto as caixas acústicas fazem o contrário, transformando os sinais elétricos em ondas sonoras audíveis.

2.1.1 Frequência

A frequência é uma das características mais importantes do áudio. O conceito de frequência é definido por Ratton (2004, p.68) como:

Um dos parâmetros fundamentais do som, caracterizando sua altura tonal. Pode ser compreendida como a quantidade de vezes que ocorre um determinado movimento periódico dentro de um intervalo de tempo. No caso específico do som, o movimento é realizado por uma fonte sonora (uma corda vibrante, o cone de um auto-falante etc.), e o intervalo de tempo é considerado como 1 segundo. Assim, a altura (tom) do som pode ser expressa a partir da frequência do movimento periódico da fonte sonora, medida em ciclos por segundo, ou Hertz, e dessa forma 1 Hertz equivale a 1 ciclo por segundo [...].

Uma vez que a frequência determina a tonalidade de um som, essa característica pode ser utilizada para diferenciação de fontes sonoras em sistemas de sonorização de múltiplas fontes. Um sistema micro processado de áudio digital poderia fazer uso dessa característica do som, por exemplo, para gerenciar e identificar oradores em uma conferência com vários participantes, uma vez que cada orador possui características tonais de voz distintas. Já em aplicações de telecomunicações que utilizam a tecnologia de voz sobre IP, é comum o uso de algoritmos de análise de frequência da voz em tempo real para obter maior eficiência e inteligibilidade na comunicação em curso.

Segundo Ratton (2007), dentro da faixa de áudio os sons podem ser classificados como:

- **Graves** – frequências abaixo de 300 Hz;
- **Médios** – frequências entre 300 Hz e 2 kHz;
- **Agudos** – frequências altas, acima de 2 kHz.

Valle (2007, p.53), por sua vez, estabelece que “A audição humana é limitada, na frequência, por um mínimo de 20Hz e por um máximo de 20kHz [...]”. Isso significa que a capacidade de ouvir cobre uma grande extensão de frequências, numa

relação de 1:1000 entre a frequência mais baixa (20Hz) e a mais alta (20.000Hz). No entanto, segundo Valle (2007, p.51), “[...] a sensibilidade do ouvido não é a mesma para todas as frequências.”, já que está diretamente relacionada à intensidade sonora e à sensibilidade de cada pessoa.

2.1.2 Amplitude

O conceito de amplitude é definido por Ratton (2004, p.16) como:

Intensidade de um sinal, normalmente percebida como volume. A amplitude de um som pode ser avaliada pela quantidade de deslocamento de ar produzido pela onda sonora. No caso do sinal de áudio num equipamento de áudio analógico (ex:amplificador), a amplitude pode ser avaliada pela quantidade de variação da tensão elétrica que representa o som. Já nos equipamentos de áudio digital, a amplitude do sinal de áudio é avaliada por dados numéricos, que representam o som digitalizado.

Ou seja, a amplitude está relacionada ao volume e determina se um som é mais fraco ou mais forte que outro, sendo que quanto maior a intensidade do som, mais energia ele contém. Assim, para que um sistema de sonorização possa produzir sons de altas intensidades é necessário utilizar amplificadores de sinais capazes de gerar grande quantidade de potência.

Nesse sentido, Ratton (2004) define também o conceito de “faixa dinâmica” como sendo a diferença entre os níveis mínimos e máximos de intensidade do sinal que um equipamento de som pode distinguir ou produzir, geralmente expressa em decibéis - dB.

Na esteira da definição anterior, Valle (2007) estabelece que a faixa dinâmica da audição humana é limitada, em intensidade, por um mínimo de 0dB e por um máximo de 120dB.

2.1.3 Decibel (dB)

Ratton (2004, p.44) define decibel (dB) como “Unidade de medida usada para avaliar o nível de potência de áudio, através de uma relação logarítmica entre dois números (o que significa que a medida é sempre em relação a uma referência) [...]”.

Nota-se que a escala logarítmica em dB foi adotada para medições de áudio em razão da grande faixa de amplitude sonora que pode ser detectada pelo ouvido humano, desde pequenos ruídos até grandes shows de rock! E, por ser logarítmica, a medida em decibel apresenta uma característica não linear, também se adequando à maneira com que o ouvido humano percebe as intensidades sonoras.

Segundo Valle (2007), uma variação de dez vezes na potência sonora soa para o ouvido humano como “o dobro do som”, motivo pelo qual se adotou a base 10 para medidas de nível sonoro. Valle (2007, p.42) destaca ainda alguns fatores que endossam a utilização do decibel como unidade de medida de som em sistemas de áudio:

- A menor diferença de nível sonoro que o ouvido humano consegue perceber é de aproximadamente 1dB.
- Uma variação para um valor maior produz um número positivo de dB, e uma variação para um valor menor produz um número negativo de dB.
- 0dB significa que não houve variação alguma.
- Se a variação for para zero de potência (sinal completamente retirado), tem-se $-\infty$ dB.

Assim como no caso das frequências, o ouvido humano também só é capaz de perceber sons que estejam dentro de determinada faixa de intensidades. A Figura 2 a seguir apresenta a escala de níveis de intensidade sonora em *dB SPL* – “*Sound Pressure Level*”, cuja referência é o valor de pressão $20 \mu\text{N/m}^2$, que corresponde ao limiar da audição (0 dB).

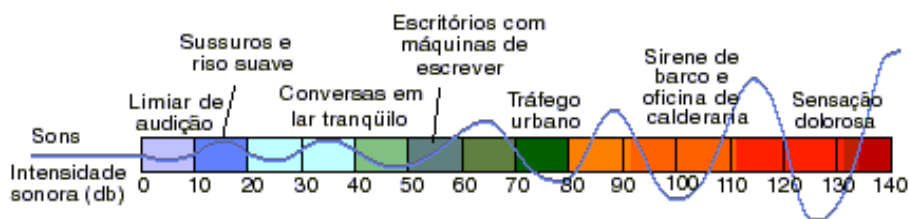


Figura 2: Escala em dB SPL.

Fonte: www.prof2000.pt (2010)

As características de amplitude e frequência do sinal de áudio se relacionam de modo que o ouvido não possui a mesma sensibilidade às intensidades sonoras em toda a faixa de frequências audíveis (20 Hz a 20 kHz). Segundo Valle (2007, p.51), “[...] o ouvido é mais sensível às frequências entre 3kHz e 5kHz, piorando fortemente em direção aos graves e piorando um pouco em direção aos agudos. [...]”. Isso significa que para ter a mesma percepção de volume em todo o espectro de frequências é necessário aplicar níveis distintos de pressão sonora, tal como evidenciado nas curvas do gráfico na Figura 3 a seguir, popularmente conhecida como *Curvas de Fletcher e Munson*. Note que a percepção das frequências varia em função da intensidade do som. Essa sensibilidade também pode ser alterada em razão do avanço na idade (envelhecimento) do ouvinte.

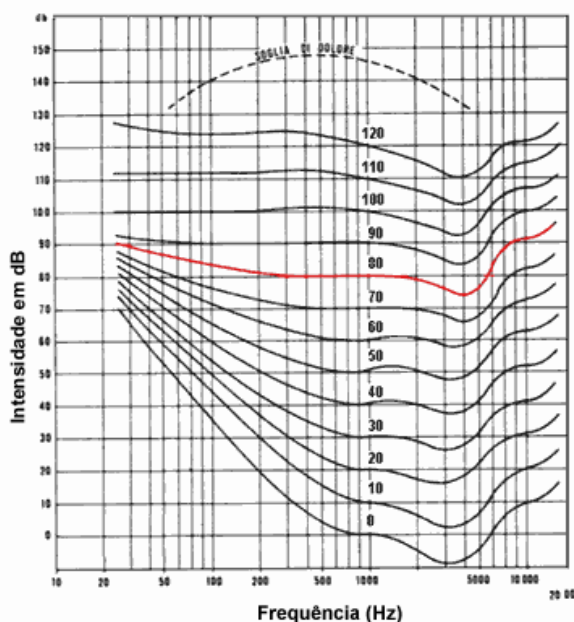


Figura 3: Curvas de sensibilidade sonora.

Fonte: <http://telecom.inescporto.pt> (2010)

2.2 ÁUDIO DIGITAL

A digitalização de sinais é uma das tecnologias mais utilizadas na atualidade, com aplicações em diversas áreas como música, telefonia, processamento de imagens, entre outros. Na área de áudio, essa forma de representação amplia consideravelmente as possibilidades de edição, armazenamento e processamento de informações sonoras em relação ao modelo analógico, sendo a base fundamental para implantação e operacionalização de um sistema de áudio sobre IP.

Ratton (2004, p.19) define Áudio Digital como “Forma de se representar a onda sonora, de tal maneira que cada ponto da onda é codificado por um número.[...]”. Em outras palavras, trata-se da representação de uma onda sonora (ou de um sinal elétrico de áudio analógico) no formato de níveis discretos, onde cada ponto da onda é associado a um número do sistema binário.

2.2.1 Digitalização do Áudio

Enquanto os sinais analógicos são de natureza contínua, isto é, possuem um valor numérico em qualquer instante de tempo, os sinais digitais (ou discretos) só possuem valor numérico em determinados instantes de tempo, onde para se determinar os valores entre dois instantes de tempo conhecidos são utilizados cálculos matemáticos de interpolação numérica.

A conversão de um sinal de áudio analógico para o formato digital é realizada utilizando um dispositivo chamado Conversor A/D, que a intervalos constantes colhe amostras do sinal analógico, como se tirasse “fotografias” do sinal de áudio, e as converte em números binários que as representam no domínio digital como uma série de valores numéricos, como ilustrado na Figura 4.

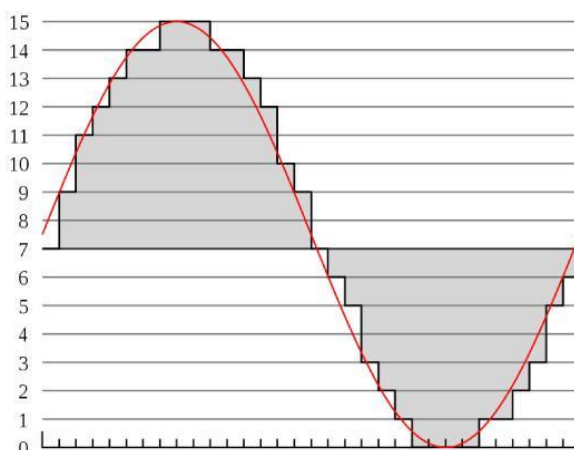


Figura 4: Digitalização do áudio.

Fonte: www.videomaker.com (2010)

Sobre o conversor A/D, Ratton (2004, p.39) traz a seguinte definição:

O conversor A/D (analógico/digital) é o dispositivo que traduz em códigos numéricos digitais as variações e flutuações contínuas de um sinal (vindo de um microfone, por exemplo). No processo de conversão de analógico para digital, o conversor A/D faz inúmeras amostragens do valor do sinal analógico ao longo do tempo [...].

No sentido oposto, um Conversor D/A reconstrói o sinal original de áudio analógico a partir dos códigos numéricos binários que o representam digitalmente, pulso a pulso, respeitando sua ordem sequencial no tempo e na mesma taxa de amostragem em que foi digitalizado.

Como o som que produzimos e ouvimos é de natureza estritamente contínua e analógica, materializado fisicamente pela vibração das moléculas de ar ao nosso redor, os conversores A/D e D/A cumprem importante função de interface entre o mundo real e os sistemas digitais, permitindo que o som se propague adequadamente nesses dois meios. Por exemplo, a voz de um orador produz ondas sonoras acústicas. Um microfone converte essas ondas sonoras em um sinal elétrico analógico contínuo. Para que se possa processar esse sinal digitalmente, é necessário convertê-lo para o domínio digital. Por fim, após processado, é necessário convertê-lo de volta ao formato contínuo no tempo para que possa ser reproduzido nos auto-falantes.

2.2.2 Taxa de Amostragem

A definição de Ratton (2004) para Taxa de Amostragem é a seguinte:

Frequência ou taxa com que um conversor A/D efetua as amostras de um som durante sua digitalização. A taxa de amostragem (sampling rate) é medida em Hertz, e deve ser igual a pelo menos o dobro da maior frequência existente no sinal.

Ou seja, diz respeito à velocidade de atuação do conversor A/D no processo de amostragem do sinal analógico. Desse modo, a taxa de amostragem deve ser alta o suficiente para capturar as rápidas variações das componentes de alta frequência do sinal de áudio.

Quando o autor diz que a taxa de amostragem “deve ser igual a pelo menos o dobro da maior frequência existente no sinal”, está se referindo ao Teorema de Nyquist, que diz que para que se capturem todas as frequências componentes de um sinal é necessário que sua amostragem ocorra a uma taxa superior ao dobro da maior frequência existente no sinal que se quer amostrar. Nesse caso, como a faixa de áudio compreende o intervalo de 20 Hz a 20 kHz, então, para se obter uma representação fiel de qualquer sinal de áudio, a taxa de amostragem deverá ser superior a 40 kHz. Essa é a razão pela qual é comumente utilizada a taxa de 44.100 amostras de áudio por segundo, ou 44.1 kHz em CDs e DVDs de música. Porém, com o significativo aumento da capacidade de processamento e armazenamento dos sistemas digitais, taxas de amostragem superiores como 96 kHz e 192 kHz tem sido utilizadas em aplicações de áudio de alta definição.

2.2.3 Resolução

De acordo com a definição de Ratton (2004, p.127), “[...] resolução é a precisão da medida do nível do sinal, quando as amostras são codificadas em números, e está diretamente relacionada ao número de bits utilizado para a representação desses

números. [...]”. Em outras palavras, pode-se dizer que o termo “Resolução” está relacionado à capacidade do conversor A/D de detectar fielmente as variações de amplitude do sinal analógico, e que quanto maior for a resolução do processo de digitalização, mais fiel será a representação do sinal original, o que implica em melhor qualidade do áudio.

Note, portanto, que a digitalização de um sinal de áudio analógico fundamenta-se essencialmente em dois processos: a amostragem de tempo e a amostragem de amplitude. No segundo caso, porém, há a ocorrência de um “efeito colateral” conhecido como Erro de Quantização, que segundo Ratton (2007), ocorre quando o valor real da amostra está entre dois níveis da escala e há o arredondamento para o nível disponível mais próximo, já que é impossível para o conversor determinar um valor intermediário. Esse processo de arredondamento gera ruído, que tende a ser reduzido com o aumento da quantidade de bits que representa cada amostra.

Pode-se dizer então que o erro de quantização ocorre porque, teoricamente, os sistemas digitais não possuem quantidades de bits ilimitadas para representar todos os valores possíveis de um sinal analógico contínuo, surgindo à necessidade de se arredondar a amostra para o valor digital mais próximo. Logo, quanto maior a resolução do conversor, maior será a precisão na digitalização do sinal de áudio analógico.

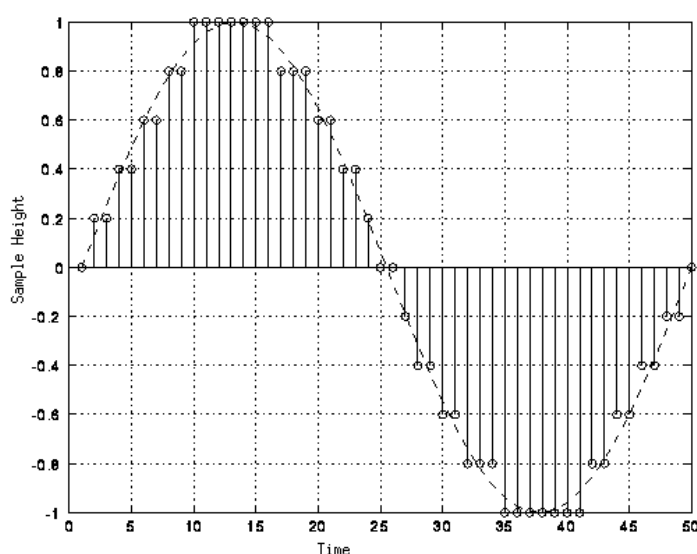


Figura 5: Resolução e Quantização de um sinal de áudio.

Fonte: www.webkinesia.com (2006)

Ratton (2007) estabelece que uma resolução de 16 bits, comumente usada nos CDs de áudio, permite a digitalização do sinal analógico em 65.536 níveis distintos de amplitude, o que representa cerca de 96 dB de faixa dinâmica, considerado adequado para a reprodução de música em geral. No entanto, cumpre acrescentar que o crescente aumento na capacidade de processamento e armazenamento de dispositivos digitais, aliado à redução de custo dos conversores A/D, já tem possibilitado a utilização de resoluções de 24 e 32 bits em aplicações de áudio de alta definição.

2.3 SISTEMAS DE SONORIZAÇÃO

Um sistema de sonorização nada mais é que um conjunto de dispositivos e equipamentos elétricos, eletrônicos, eletromecânicos e digitais, interligados ordenadamente de modo a viabilizar a captação, processamento, armazenamento e distribuição do som originado de uma ou mais fontes sonoras, em ambientes abertos ou fechados de grande ou pequeno porte. É utilizado nas mais variadas aplicações, tais como auditórios, salas de concerto, salas de conferência, estúdios de gravação, estádios, igrejas, shopping centers, discotecas, shows de música, entre outros.

Tradicionalmente, um sistema de sonorização é subdividido da seguinte maneira:

- **Fonte:** gerador das ondas sonoras, tal como um orador ou um instrumento musical;
- **Captação:** transformação das ondas sonoras em sinais de áudio (analógico ou digital). O microfone é um tipo de dispositivo que executa essa função;
- **Equalização:** alteração das características dos sinais de áudio através da atenuação ou realce de determinadas frequências. Os controles de graves, médios e agudos de um equalizador são um exemplo típico.
- **Processamento:** manipulação do sinal de áudio para alterações de ganho, fase, tempo, frequência, distorção, etc. Dispositivos como compressores, gates e reverbs são exemplos clássicos de processadores de sinais;

- **Mixagem:** mistura ou somatória dos diversos sinais de áudio provenientes das diferentes fontes sonoras para posterior direcionamento e encaminhamento do sinal resultante a outros dispositivos. O equipamento que realiza essa função é popularmente conhecido como mixer, mesa de som ou mesa de mixagem;
- **Gravação:** registro dos sinais de áudio em mídias de armazenamento (CD, DVD, fita magnética, disco rígido, etc) para posterior recuperação das informações sonoras gravadas;
- **Transmissão:** distribuição dos sinais de áudio para receptores remotos localizados em locais distintos ao da geração dos sinais, como é o caso das emissoras de rádio, televisão e internet;
- **Amplificação:** aumento da amplitude do sinal de áudio para um nível de potência suficiente para reprodução nos autôfalantes;
- **Reprodução:** emissão do som no ambiente a partir da transformação dos sinais de áudio em ondas sonoras através da vibração dos autôfalantes.

A Figura 6, na página seguinte, apresenta o diagrama esquemático de um sistema tradicional de sonorização e imagem analógico geralmente encontrado em salas de conferência e auditórios. Já a Figura 7 ilustra a configuração mais usual de sistemas de sonorização utilizados em shows e eventos musicais.

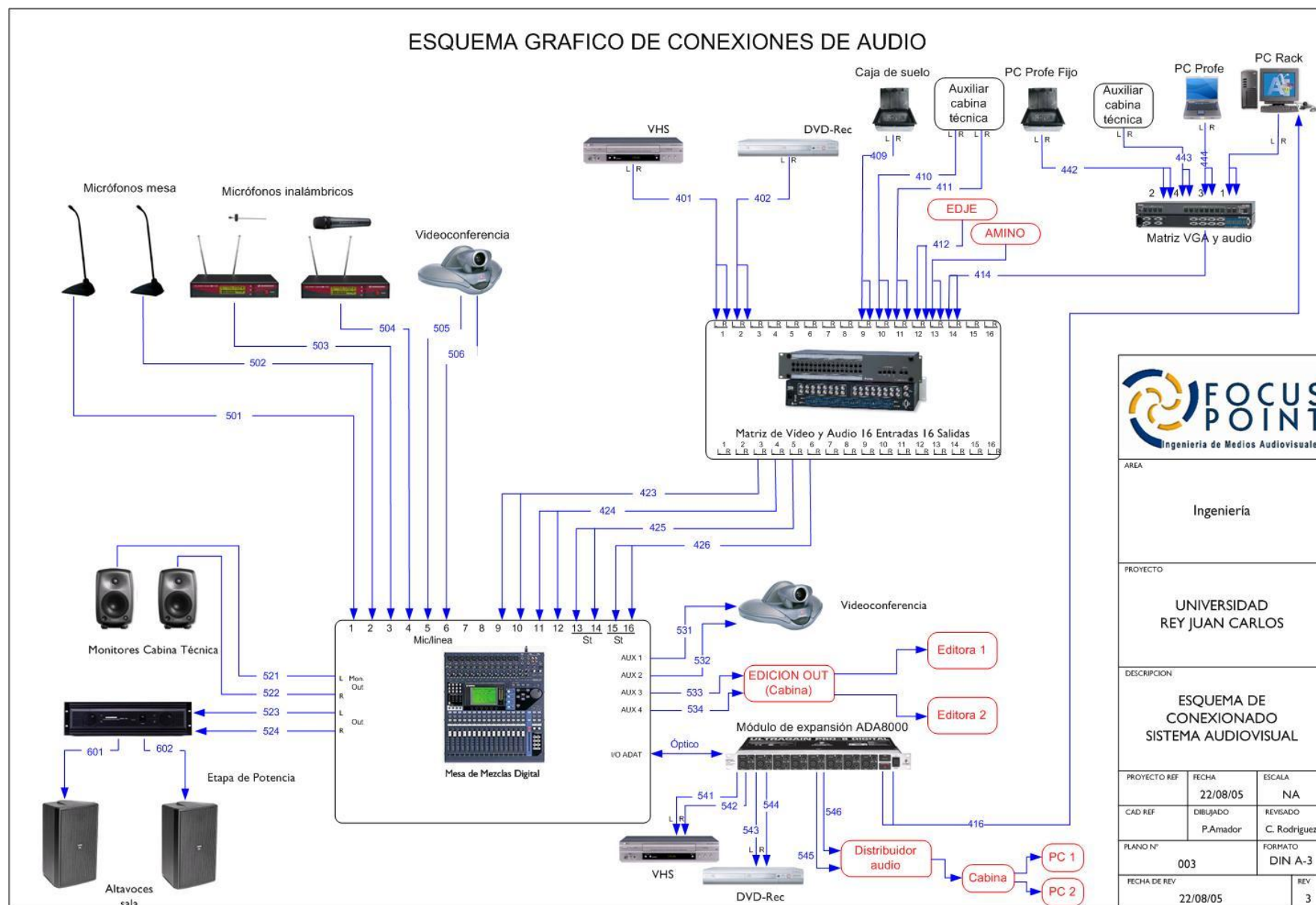


Figura 6: Exemplo típico de um sistema analógico de sonorização e imagem para salas e auditórios.

Fonte: Focus Point (2005), <http://chaplin.urjc.es>.

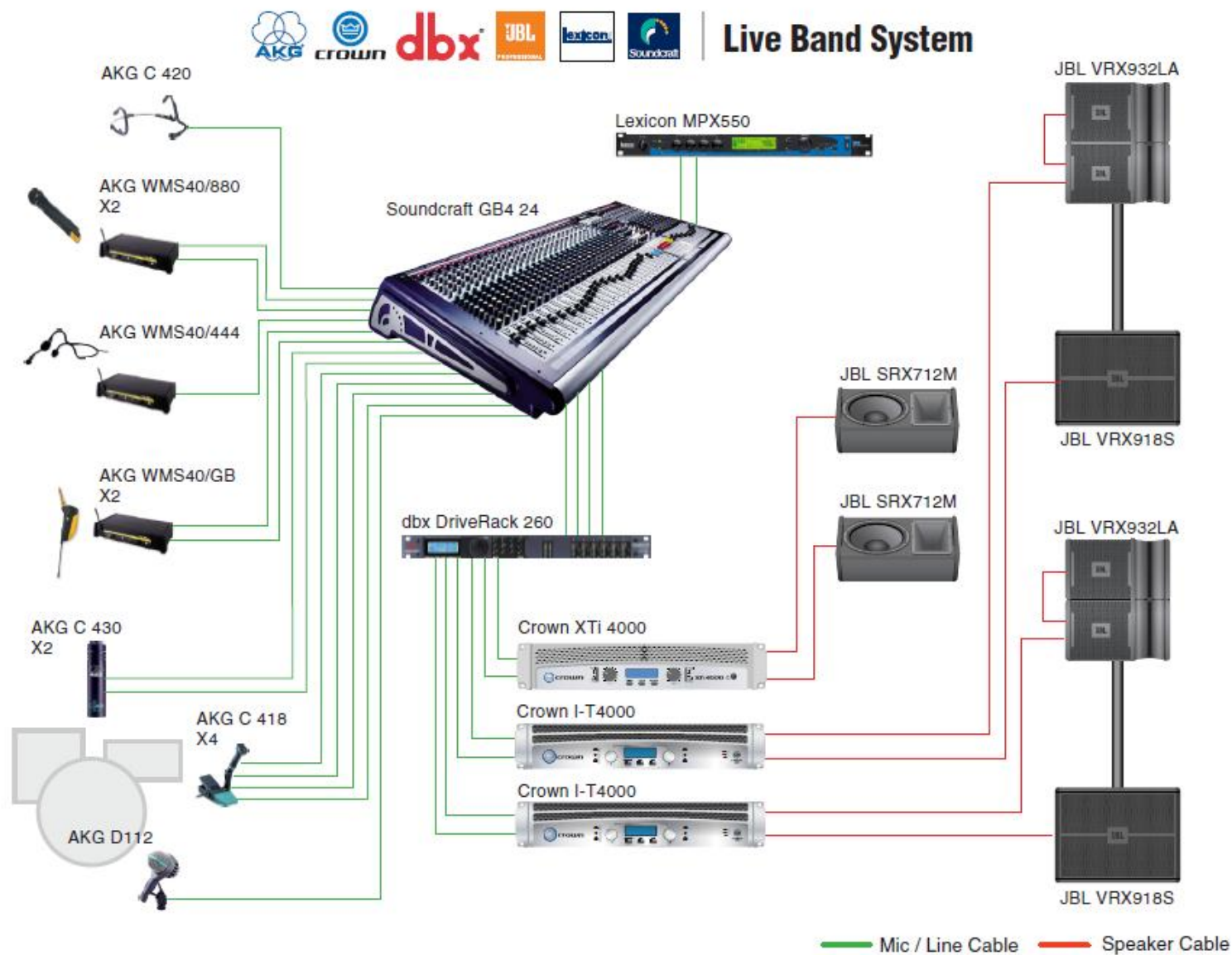


Figura 7: Exemplo típico de um sistema analógico de sonorização musical.

Fonte: Crown Audio (2010), www.crownaudio.com.

3 REDES ETHERNET/IP

As redes de dados Ethernet/IP tornaram-se um padrão mundial para comunicação de dispositivos em redes de telecomunicações, apresentando características de grande confiabilidade, eficiência, flexibilidade, escalabilidade e baixo custo. Sua estrutura baseada em camadas permite que dispositivos das mais variadas funções e de diferentes fabricantes possam se comunicar entre si, na mesma rede ou entre redes distintas, dado seu alto grau de interoperabilidade e padrão uniforme de endereçamento.

São baseadas em um conjunto de hardware e software que viabilizam a troca de informações e o compartilhamento de recursos entre os dispositivos de uma rede local (LAN), tais como servidores, estações de trabalho, roteadores, switches, modems, impressoras, firewalls, sistemas operacionais, bancos de dados, entre outros. As informações trafegam em alta velocidade através de um único cabo e alcançam todos os nós endereçáveis na rede.

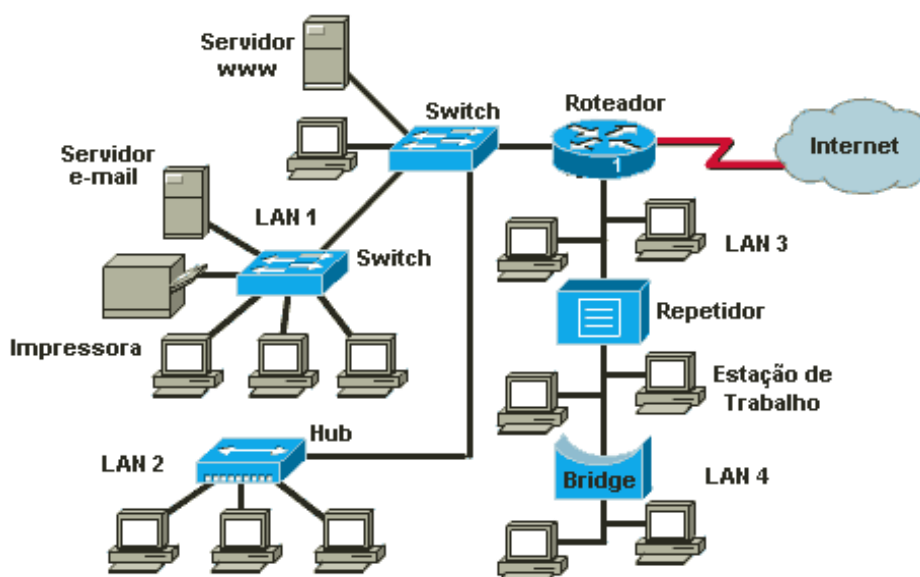


Figura 8: Exemplo típico de uma rede Ethernet/IP.

Fonte: <http://3.bp.blogspot.com> (2010)

3.1 MODELO OSI

O Modelo OSI - *Open Systems International* – é a referência de estruturação das redes de telecomunicações baseadas em pacotes, criado com a finalidade de padronizar o desenvolvimento de equipamentos para redes de comunicação de dados. Trata-se de um conjunto de sete camadas que define os diferentes estágios pelos quais os dados devem trafegar de um dispositivo para outro dentro de uma rede.

O padrão foi estabelecido pelos órgãos *International Standards Organization* (ISO) e *International Telecommunications Union* (ITU-T). As três camadas superiores são especificadas e documentadas pelo *Internet Engineering Task Force* (IETF), enquanto as demais são normatizadas pelo *Institute of Electrical and Electronics Engineers* (IEEE).

Tabela 1: Modelo OSI em camadas.

Camada	Nome	Função
7	Aplicação	Fornece serviços às aplicações do utilizador.
6	Apresentação	Encriptação e compressão de dados. Assegura a compatibilidade entre camadas de aplicação de sistemas diferentes.
5	Sessão	Controla as sessões entre aplicações.
4	Transporte	Controle de fluxo de informação, segmentação e controle de erros.
3	Rede	Encaminhamento de pacotes e fragmentação. Esquema de endereçamento lógico.
2	Dados	Controla o acesso ao meio físico de transmissão. Controlo de erros da camada física.
1	Física	Define as características do meio físico de transmissão da rede, conectores, interfaces, codificação ou modulação de sinais.

3.2 PROTOCOLO IP

O Protocolo IP (*Internet Protocol*) fornece um mecanismo de endereçamento uniforme que possibilita a comunicação entre dispositivos de uma rede ou de redes distintas, onde cada dispositivo possui um endereço único de identificação que o permite enviar e receber pacotes de dados.

IP é o formato padrão para transmissão de qualquer tipo de dado na internet ou em redes locais, cabeadas ou sem fio, onde todo e qualquer computador produzido hoje em dia vem equipado com um adaptador de rede IP. Além disso, possui arquitetura aberta e independente, dispensando qualquer tipo de interface ou aplicação proprietária.

Com a popularização e a consolidação do protocolo IP como padrão para comunicação de dados, outras indústrias passaram a adotá-lo na modernização de seus sistemas e produtos, tornando sua utilização não mais restrita às redes de computadores. Um exemplo disso são as empresas de telefonia que estão abandonando seus tradicionais sistemas chaveados e os substituindo pelo modelo de voz sobre IP (VoIP – Voice Over IP), tanto em aplicações corporativas quanto em serviços pessoais para o consumidor final. Outro exemplo são as emissoras de rádio e televisão que, após a migração para os sistemas digitais, demonstram uma tendência de utilização de redes IP para distribuição de seu conteúdo multimídia. Essa tecnologia tem sido incorporada e utilizada até mesmo em produtos domésticos como aparelhos de som, TV, porta-retratos e cafeteiras, inserindo-se no conceito de Casa Inteligente em aplicações de automação residencial.

Segundo Simpson e Greenfield (2007), um dos motivos que contribuem para a popularização da tecnologia IP é o seu baixo custo de hardware. Um adaptador de rede Gigabit Ethernet, por exemplo, que opera a taxa de 1000 Mbps, é encontrado no mercado de peças e componentes de informática ao preço médio de \$15 (dólares americanos), e o preço cai continuamente! Outro motivo favorável, segundo os autores, diz respeito à capacidade dos sistemas IP de suportar a crescente demanda dos usuários por qualidade e disponibilidade das redes, o que impõe a

necessidade de maiores velocidades e largura de banda. Isso vem ocorrendo a cada ano que se passa nas últimas três décadas e a tecnologia IP sempre se mostrou preparada e capaz de atender e acompanhar essa evolução.

Enfim, assim como o computador pessoal tornou-se a plataforma padrão para as mais diversas aplicações pessoais e profissionais, com velocidade, poder de processamento e custo altamente benéficos, o protocolo IP tem se tornado o formato padrão para transporte de quaisquer tipos de dados.

4 ÁUDIO SOBRE IP

Os sistemas de áudio profissional, cuja topologia tradicional foi apresentada anteriormente na Figura 6, tem sido objeto de profundas modificações desde que as redes IP passaram a ser incorporadas em sua estrutura, promovendo uma revolução na maneira como os sinais de áudio são gerenciados e distribuídos e proporcionando vantagens e possibilidades até então inexploradas.

O método de funcionamento de um sistema de áudio sobre IP se assemelha a uma tradicional rede de comunicação de dados entre computadores e periféricos, tal como resumido por Novak (Axia Audio):

- Fontes de áudio individuais (microfones, instrumentos musicais, etc) são conectadas a “nós de áudio” (Figura 9), que convertem os sinais analógicos em sinais digitais a taxas acima de 44 kHz de amostragem e 24 bits de resolução.
- Uma vez digitalizado, o áudio é codificado e encapsulado em pacotes de dados no formato adequado para transmissão na rede;
- Para cada entrada e saída dos nós de áudio é atribuído um endereço IP único, para fins de identificação e encaminhamento;
- Dispositivos de áudio com funções de sincronia e sinalização também são conectados aos nós de rede para conversão dessas informações de controle em pacotes de dados;
- Cada nó é então conectado às portas de um switch local que disponibiliza o áudio e os dados de controle na rede de alta capacidade, onde links Gigabit Ethernet ou fibra ótica permitem a transmissão de milhares de sinais de áudio e dados simultâneos;

Assim, segundo Novak, o conteúdo e controle de todos os dispositivos de áudio passam a estar disponíveis a partir de qualquer local na rede, e como essas informações foram codificadas no formato de dados, é possível gerenciar toda a transmissão por meio de softwares automatizados através de um computador.

4.1 COMPONENTES DO SISTEMA

A composição de um sistema de áudio sobre IP pode variar de acordo com a sua complexidade e aplicação, mas, na maioria dos casos, fazem parte do arranjo os seguintes itens:

- **Fontes de Áudio:** assim como nos sistemas analógicos, compreende os microfones e instrumentos para captação e/ou geração das ondas sonoras;
- **Nós de Áudio:** equipamentos localizados próximos às fontes de áudio com a função de fazer a interface com a rede de comunicação. Convertem, digitalizam, comprimem e endereçam os sinais conectados nas entradas e fornecem como saída o áudio codificado e encapsulado no formato de dados padrão da rede. Geralmente possuem display e leds de monitoração frontais e aplicativo de gerenciamento embarcado para funções de programação, configuração, monitoramento e comando, com acesso local ou via rede através de um PC. A Figura 9 mostra um exemplo desse tipo de equipamento com capacidade para receber oito microfones.



Figura 9: "Nó de Áudio" analógico ip88m Analog Mic I/O Blade.

Fonte: Wheatstone Corporation (2009), www.wheatstone.com.

- **Engine:** hardware central de processamento dos sinais de áudio da rede. Executa funções de mixagem, equalização, dinâmica, efeitos, etc, substituindo o mixer e os periféricos dos sistemas analógicos tradicionais;

- **Console:** plataforma de comando e controle que agrega as mesmas funções tradicionais das mesas de som analógicas e digitais. Dispõe de faders, slides, knobs, leds, display e botões com atuação e monitoração em tempo real na rede, tal como apresentado na Figura 10;



Figura 10: Console Axia Element.

Fonte: Axia Audio (2007), www.axiaaudio.com.

- **PCs e Servidores:** computadores com funções de acesso, gerenciamento, armazenamento, automação, controle, registros, backups, redundância, entre outros;
- **Softwares:** aplicativos que rodam nos PCs e servidores com a função de prover uma interface gráfica entre os usuários e a rede para fins de monitoração, programação e gerenciamento. A Figura 11 ilustra um exemplo desse tipo de aplicação.

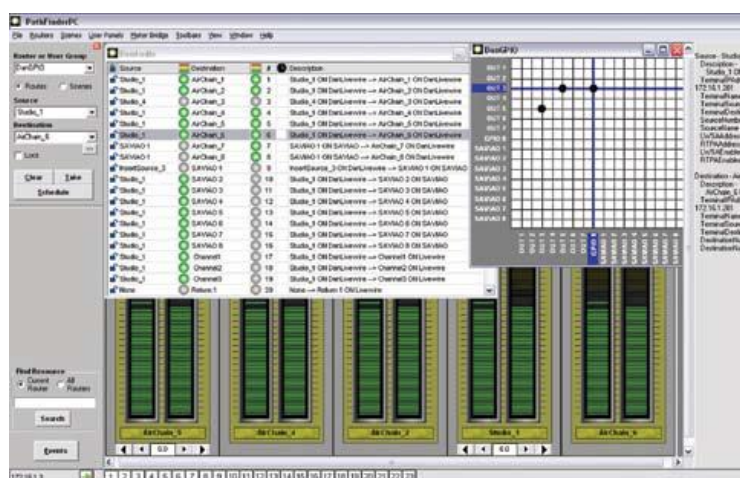


Figura 11: Software de gerenciamento PathfinderPC.

Fonte: Axia Audio (2007), www.axiaaudio.com.

- **Dispositivos de Monitoração de Áudio:** tal como nos sistemas analógicos, são responsáveis pela projeção do som no ambiente ao converter os sinais de áudio em ondas sonoras. Compreendem os autôfalantes, fones de ouvido e caixas de som geralmente conectados a nós de áudio;
- **Switches/Roteadores:** possuem a função de interconectar todos os dispositivos da rede controlando a comunicação entre eles. Geralmente são utilizados switches do tipo “gerenciados” que permitem configuração e monitoração de parâmetros internos, estabelecimento de prioridades e bloqueios, entre outras funcionalidades.

A Figura 12 ilustra um exemplo típico de um sistema de áudio sobre IP que inclui os principais elementos anteriormente mencionados.

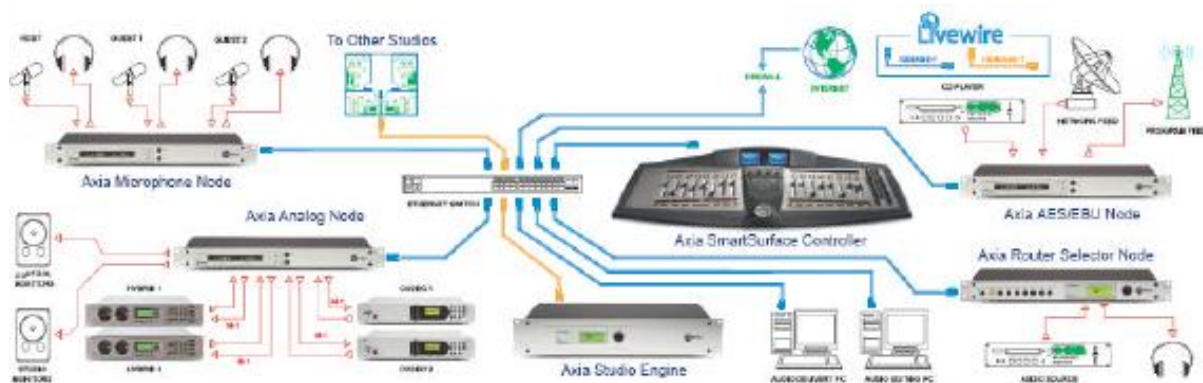


Figura 12: Esquema típico de um sistema de áudio sobre IP.

Fonte: Axia Audio (2007), www.axiaaudio.com.

4.2 ARGUMENTOS PARA ÁUDIO SOBRE IP

Pizzi e Church (2010) elencam os principais argumentos que favorecem a implantação de um sistema de áudio sobre IP. São eles:

- **Escalabilidade:** talvez a maior vantagem em relação aos sistemas tradicionais analógicos é a facilidade de ampliação e mudanças em sistemas de arquitetura Ethernet/IP;
- **Conveniência:** documentação enxuta, facilidade de reconfiguração e rápida instalação fazem dos sistemas de áudio sobre IP uma opção de extrema simplicidade às áreas de engenharia e operação;
- **Integração:** inúmeros benefícios e melhorias podem ser obtidos por meio da integração e interconexão com outras plataformas também baseadas em IP, tais como sistemas telefônicos de voz sobre IP, sistemas de segurança e supervisão predial, sistemas de detecção e combate a incêndios, entre outros;
- **Linguagem dos PC:** atualmente a grande maioria dos equipamentos de áudio profissional possui algum tipo de driver, aplicação ou software de gerenciamento e controle que opera a partir de um computador local conectado a uma porta serial USB, Firewire, RS-232 ou similar. Como o protocolo IP já é por natureza a linguagem padrão dos computadores, um sistema de áudio sobre IP permite se conectar a um ou mais computadores por meio de um único cabo com terminação RJ-45, propiciando uma poderosa interface de baixo custo e alta resolução;
- **Popularidade:** existe uma grande quantidade de informações e recursos disponíveis, tais como livros, revistas, sites e cursos especializados que abordam as tecnologias de redes IP, assim como equipamentos, cabos conectores, ferramentas e mão-de-obra relacionados podem ser facilmente encontrados no mercado;

- **Preparado para o futuro:** uma arquitetura de áudio baseada em IP oferece grande flexibilidade, o que possibilita a realização de mudanças significativas em todo o sistema para fins de ampliação ou modernização sem grandes transtornos e a custos relativamente baixos, se comparado aos modelos tradicionais. Além disso, a adoção do padrão IP em sistemas de áudio profissional permite a aplicação da Lei de Moore, que estabelece que a capacidade e poder dos equipamentos dobra a cada dois anos;
- **Custo:** um sistema de áudio sobre IP tende a ser mais econômico que um sistema tradicional, onde a diferença de custos se amplia na medida em que cresce o tamanho das instalações (vide tópico 4.6).
- **Padronização:** entidades internacionais estabeleceram um padrão de requisitos e formalidades relacionados aos sistemas de áudio sobre IP que garantem a interoperabilidade entre os equipamentos de diferentes fabricantes, tal como descrito no tópico 4.8 a seguir.

4.3 INFRAESTRUTURA DE REDE

Os sistemas de áudio sobre IP requerem redes de telecomunicações bem gerenciadas, com boa qualidade de serviço (QoS) e larga banda de transmissão alocada. Em razão de tais requisitos é que esses sistemas são projetados predominantemente para operar em infraestruturas de redes locais (LAN) confiáveis e de baixa latência, adequadas para aplicações profissionais. Também recomenda-se que a rede seja de utilização exclusiva para transmissão e distribuição de áudio, em infraestrutura independente das redes de dados já existentes no local.

De acordo com Wheatstone Corporation (2009), a especificação e implementação de redes de telecomunicações destinadas a aplicações de áudio profissional deve considerar os seguintes parâmetros:

- **Escopo:** estabelecer as necessidades atuais e provisionar as necessidades futuras;
- **Infraestrutura física:** definir a topologia de rede, tipo de cabeamento (CAT-5e/CAT-6, fibra ótica), alocação de switches e roteadores;
- **Taxa de Transferência:** verificar se a taxa de transferência (ou *throughput*) dos roteadores e switches é suficiente;
- **Capacidade:** conferir se há espaço/banda para ampliação da rede;
- **Aplicação:** definir se a rede será usada exclusivamente para aplicações de áudio ou se será compartilhada com outras aplicações;
- **Atualizações:** considerar meios para atualizações e *upgrades*;
- **Monitoração:** identificar as ferramentas de software necessárias para fins de monitoração e controle;
- **Acesso Remoto:** disponibilizar um acesso seguro para monitoração remota e suporte técnico à rede.

4.4 STREAMING X IP

Segundo Pizzi e Church (2010), quando o áudio deixa o ambiente seguro e controlado das redes locais dedicadas e passa a ser transmitido em redes públicas, sua denominação passa a ser “*streaming de áudio*”, não mais áudio sobre IP. Os autores defendem que, apesar de estarem relacionados, esses conceitos não se confundem. O streaming é utilizado predominantemente para distribuição de conteúdo multimídia na internet a altas taxas de compressão que comprometem a qualidade do material transmitido, sem garantias de entrega e com atrasos da ordem de vários segundos, a depender da disponibilidade e utilização das redes. Como exemplo é citado o popular site You Tube – www.youtube.com.

Além disso, Tieline (2010) menciona que os Provedores de Serviço de Internet (ISP) em geral bloqueiam o tráfego de pacotes de dados originados de transmissões *multicast*, método predominante nas aplicações de áudio sobre IP, restringindo portanto seu alcance às redes locais privadas.

No entanto, apesar de se tratar de uma rede pública, não-gerenciada, de longa distância e com alto tráfego de outros tipos de dados, a crescente ampliação da largura de banda e serviços disponíveis na internet deve levar a sua efetiva utilização num futuro próximo, em uma provável segunda geração de sistemas de áudio sobre IP.

4.5 VOZ SOBRE IP X ÁUDIO SOBRE IP

Tesch (2007) faz uma importante diferenciação entre os sistemas de voz e áudio sobre IP, que apesar de estarem relacionados, possuem características e aplicações bastante distintas.

A tecnologia de voz sobre IP (VoIP) é predominantemente utilizada para chamadas telefônicas de duração relativamente curta, com baixa taxa de amostragem do áudio captado, geralmente 8 kHz, bem abaixo do mínimo de 44 kHz exigido nas aplicações de áudio profissional. Isso porque no caso de VoIP a informação a ser transmitida é essencialmente a voz humana, que possui particularidades psicoacústicas favoráveis à utilização de artifícios técnicos de compressão, mascaramento e redundância para obtenção de resultados satisfatórios mesmo a baixas taxas de amostragem e resolução.

Já as aplicações de áudio sobre IP caracterizam-se por longos períodos de duração e alta qualidade do material sonoro, principalmente em distribuição de conteúdo musical para o público em geral. Neste caso, qualquer tentativa de se aplicar as técnicas de VoIP no áudio transmitido causaria evidente percepção de baixa qualidade sonora por parte dos ouvintes, com sensível degradação nas regiões extremas da faixa de frequência audível.

4.6 INSTALAÇÃO E CUSTOS

Os sistemas tradicionais de áudio analógico se caracterizam por uma instalação baseada em conexões ponto a ponto e centralizadas na mesa de mixagem, na qual todas as fontes e destinações de áudio são cabeadas para ou a partir desse equipamento. Instalações desse tipo tendem a assumir grande complexidade na medida em que cresce o tamanho das instalações, gerando altos custos de instalação, dificuldade de manutenção e documentação extensa. A Figura 13 é um exemplo do que pode acontecer quando essas instalações são cabeadas sem uma metodologia adequada às necessidades de operação, manutenção e expansão do sistema.

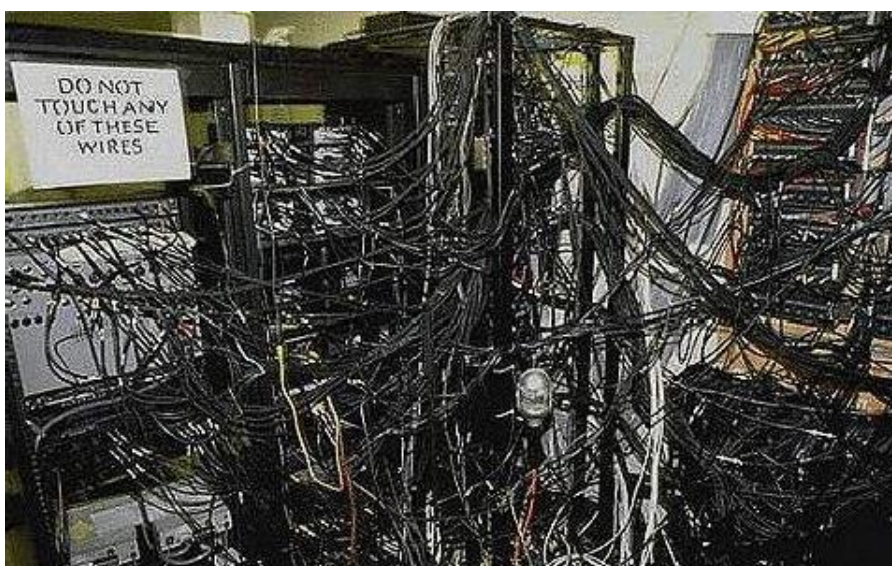


Figura 13: Cabeamento ponto a ponto mal gerenciado.

Fonte: Axia Audio (2010), www.axiaaudio.com.

Por outro lado, um sistema de áudio sobre IP tende a ser mais econômico que um sistema tradicional, seja em termos de equipamentos, materiais, instalação, operação e manutenção. A simples redução na quantidade de cabos e conectores já é suficiente para proporcionar grande economia. As tabelas seguintes apresentam dados comparativos reais dos custos (em dólares americanos) de infraestrutura e instalação de sistemas de áudio analógico e IP para ambientes típicos de estúdios de gravação.

Tabela 2: Custos de infraestrutura, equipamentos e cabeamento.

Materiais	Analógico (US \$)	IP (US \$)
Cabo CAT-6 ou Fibra Ótica	\$0	\$600
Cabo de áudio multitar	\$2.800	\$0
Guias de cabos e blocos de conectorização	\$1.600	\$0
Amplificadores de distribuição	\$2.400	\$0
Distribuidor central de áudio	\$60.000	\$18.000
Nós de áudio	\$0	\$32.300
Mesa de mixagem	\$76.000	\$68.000
Conectorização	\$900	\$1.200
Custo Total	\$143.700	\$120.100

Fonte: Audio over IP, Steve Church & Skip Pizzi (2010)

Tabela 3: Custos de instalação e mão-de-obra.

Tarefa	Analógico (horas)	IP (horas)
Interconexão de equipamentos e periféricos no estúdio	96	32
Conexão à painéis multicabos no estúdio	32	0
Lançamento de cabos multitar e terminações	48	0
Lançamento de cabos CAT-6 e terminações	0	16
Distribuição para painéis multicabos a partir da central	32	0
Interconexão de equipamentos e periféricos na central de controle	24	4
Distribuição do sinal amplificado	8	0
Programação	4	16
Total de horas de trabalho	244	68
Custo Total (\$50/hora)	\$12.200	\$3.400

Fonte: Audio over IP, Steve Church & Skip Pizzi (2010)

O posicionamento dos nós de áudio próximos às fontes (microfones, instrumentos) e destinações (caixas de som, auto-falantes) é um dos aspectos que mais contribuem para a redução significativa na quantidade de cabos, uma vez que foi eliminada a necessidade de se conectar individualmente cada dispositivo de áudio à mesa de mixagem central, muitas vezes localizada a grandes distâncias. Segundo a empresa Axia Audio (2010), outras vantagens desse novo modelo de instalação incluem:

- praticamente todas as conexões de áudio podem ser feitas sem a utilização de solda;
- redução na variedade de tipos de terminações, pois conectores tradicionalmente utilizados como os padrões P10, P2, XLR, RCA, etc, são substituídos pelo conector de rede RJ-45;
- substituição de diversos cabos de áudio analógico pesados, espessos e caros por um único cabo de rede Ethernet categoria CAT-5e ou CAT-6;
- elimina-se o uso de equipamentos tradicionais como amplificadores de distribuição, patch bays, blocos de conectorização, cabos multipar, guias de cabo, eletrocalhas, espaço físico, entre outros não mais necessários;
- a maioria das conexões entre equipamentos é *plug and play*;
- é dispensado o uso de placas e interfaces de captura de áudio nos computadores de gravação e playback, pois conectam-se através da placa de rede Ethernet;
- o roteamento dos sinais e a associação entre os dispositivos de áudio deixam de ser realizados por cabos e conectores ponto a ponto e passam a ser configuradas por uma interface gráfica de um aplicativo web.

É por essas e outras razões que os sistemas de áudio sobre IP têm ganhado espaço e aceitação no mercado e entre os profissionais da área, ocasionando a migração dos sistemas já existentes e o surgimento de novas instalações baseadas nesse modelo. É uma tecnologia que já se encontra suficientemente consolidada para permitir uma transição tranquila e confiável das linhas tradicionais de áudio analógico para as redes de arquitetura IP, proporcionando ganhos substanciais e se adequando às necessidades futuras.

4.7 TOPOLOGIAS E APLICAÇÕES

As redes de áudio sobre IP podem ser implementadas de várias maneiras distintas, partindo de configurações ponto a ponto de extrema simplicidade até complexos sistemas redundantes e autônomos compostos por dezenas de roteadores, switches e centenas de portas e canais de áudio. Naturalmente, na medida em que se amplia a complexidade da rede, ampliam-se também os custos de sua implementação.

4.7.1 Snake

De acordo com Pizzi e Church (2010), a configuração mais simples para um sistema de áudio sobre IP é denominada “Snake”, na qual dois nós são conectados por meio de um cabo *crossover*, tal como ilustrado na Figura 14, podendo-se adicionar um switch para monitoração e configuração através de um computador. Neste caso, o sistema se apresenta como uma alternativa às tradicionais interfaces PCI e USB de captura de áudio para PC, tornando-se uma plataforma de rede para pequenos estúdios de gravação.

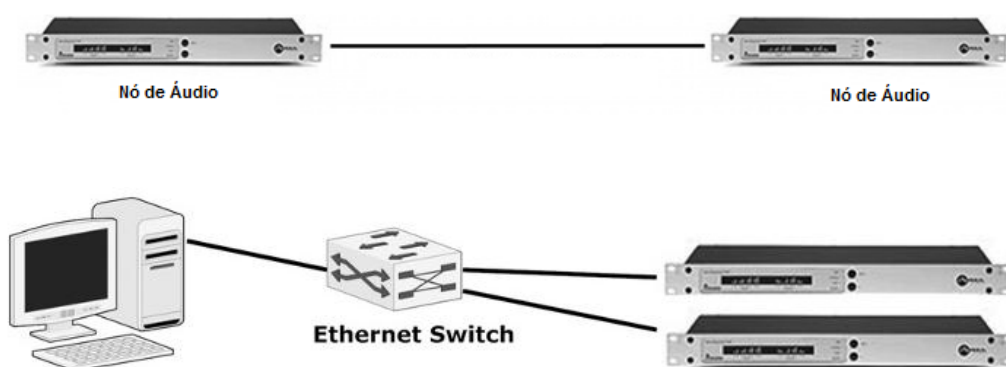


Figura 14: Exemplo de configuração "Snake".

Fonte: Audio over IP, Steve Church & Skip Pizzi (2010).

4.7.2 Studio-to-Transmitter Link (STL)

Nesta configuração, bastante similar à anterior, utiliza-se um link bidirecional de rádio Ethernet como meio de transmissão, conforme ilustração da Figura 15 a seguir.

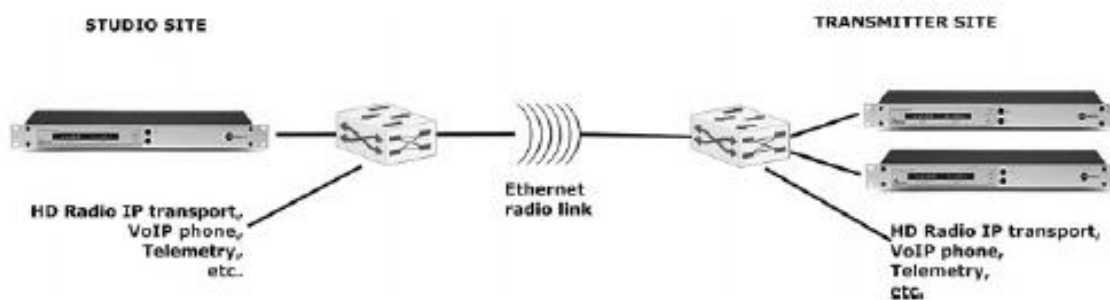


Figura 15: Configuração Studio-to-Transmitter.

Fonte: Audio over IP, Steve Church & Skip Pizzi (2010).

4.7.3 Estrela - Pequena Escala

Segundo Wheatstone Corporation (2009), essa é a configuração mais usual e recomendada para um sistema de áudio sobre IP. Consiste basicamente em um núcleo central (*Core Switch*) com funções de processamento, controle e roteamento, centralizando a conexão dos demais dispositivos periféricos que compõem a rede (Figura 16).

Apresenta como principais vantagens a simplicidade, escalabilidade e melhor desempenho, já que otimiza o tráfego dos pacotes de áudio na rede. Está limitada apenas pelo número de portas e pela taxa de transferência (*throughput*) do *Core Switch*. Por outro lado, sua grande desvantagem está na dependência crítica dos dispositivos da rede em relação ao núcleo central, já que uma falha nesse equipamento inviabiliza a operação de todo o sistema.

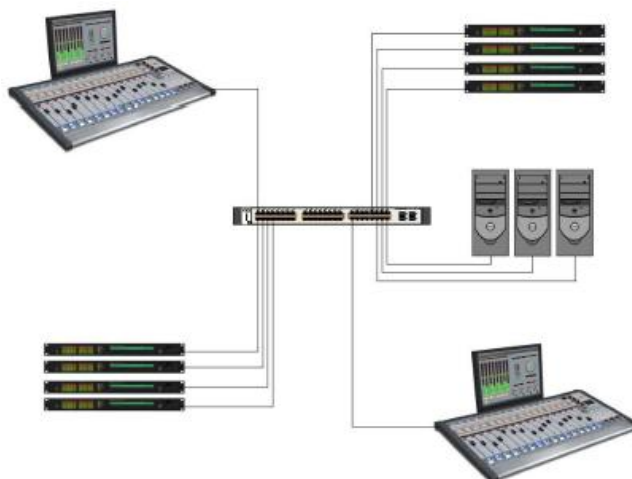


Figura 16: Topologia Estrela - Pequena Escala.

Fonte: Wheatstone Corporation (2009), www.wheatstone.com.

4.7.4 Daisy-Chained

Nessa topologia não há a presença de um switch central, onde cada subsistema possui seu próprio switch independente e interligado com os demais. É uma configuração que possui grande liberdade de expansão e possibilita o compartilhamento dos dispositivos de áudio em toda a rede, permitindo o fluxo de sinais entre seus subsistemas, assim como no caso anterior.

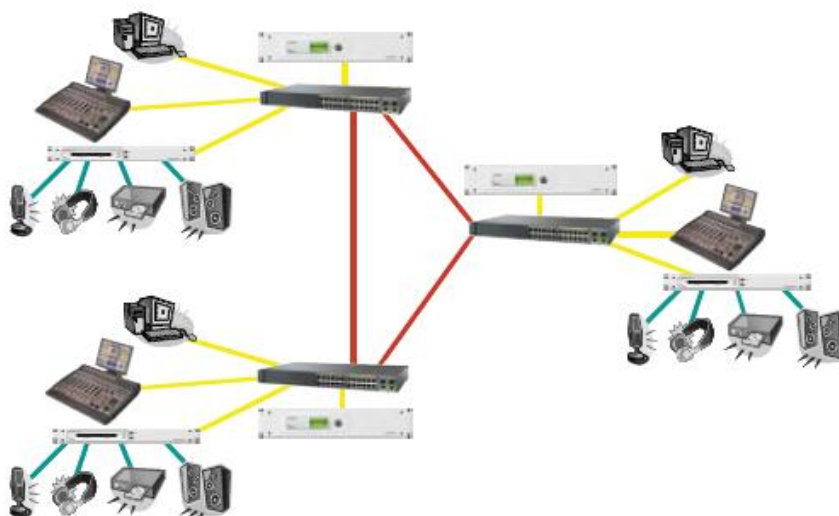


Figura 17: Configuração Daisy-Chained.

Fonte: Axia Audio (2007), www.axiaaudio.com.

4.8 PADRONIZAÇÃO EBU N/ACIP

A entidade internacional European Broadcast Union (EBU), em conjunto com fabricantes de equipamentos e profissionais da área, reuniram-se em um projeto denominado N/ACIP – Norm/Audio Contribution Over IP – cujo objetivo era estabelecer um padrão de interoperabilidade entre equipamentos de fabricantes distintos. O resultado dessa iniciativa foi o documento “*EBU Tech 3326 - Requirements for Interoperability*”, publicado em abril de 2008, bem como o documento “*EBU Tech 3329 – A Tutorial on Audio Contribution over IP*”, publicado em maio de 2008.

Tais documentos estabelecem um conjunto de requisitos mínimos necessários à compatibilidade entre dispositivos destinados a aplicações de áudio sobre IP, independentemente do fabricante. A padronização especifica os seguintes itens:

- **Protocolos de Transporte:** camada de transporte do modelo OSI, incluindo definição de portas e mecanismos de recuperação de pacotes perdidos;
- **Algoritmos de Codificação:** categorias de codecs de áudio obrigatórios, recomendados e opcionais;
- **Encapsulamento:** definições de estrutura, organização e encapsulamento do áudio em pacotes de dados IP;
- **Sinalização:** procedimentos e parâmetros de comunicação entre transmissores e receptores.

4.8.1 Protocolos de Transporte

De acordo com Harte (2007), os pacotes de dados contêm informações de endereçamento e controle que possibilitam seu tráfego nas redes de comunicação. Enquanto os dados de endereçamento (ou endereço IP) definem a origem e o

destino dos pacotes, as informações de controle determinam a maneira como serão processados ao chegarem a seu destino.

Os Protocolos de Transporte têm a função de controlar a transmissão dos pacotes de dados em conjunto com o Protocolo IP. Nesse sentido, a especificação EBU Tech 3326 adotou como padrão a utilização obrigatória dos protocolos RTP (*Real Time Protocol*) e UDP (*User Datagram Protocol*) para transporte de áudio, sendo o protocolo TCP (*Transmission Control Protocol*) de uso opcional. Harte (2007) faz uma definição sucinta desses três protocolos:

- **UDP – User Datagram Protocol:** protocolo que controla a transmissão unidirecional em uma rede de dados. Coordena a divisão dos blocos de dados em pacotes e adiciona informações de sequenciamento aos pacotes transmitidos durante uma sessão de comunicação baseada em endereçamento IP. Dessa maneira, o receptor torna-se apto a receber e reordenar os pacotes, reconstruindo os blocos de dados no seu formato original. Se comparado a outros protocolos de transporte, o UDP adiciona uma pequena quantidade de informações (dados de controle) ao cabeçalho de cada pacote, no entanto, não fornece nenhuma garantia de entrega dos dados, deixando ao usuário a responsabilidade de gerenciar e recuperar os pacotes perdidos.
- **TCP – Transmission Control Protocol:** coordena a transmissão, recepção, e retransmissão de pacotes em uma rede para garantir uma comunicação confiável, contabilizando e rastreando cada byte de dados transmitido. Requer que a conexão entre transmissor e receptor seja previamente estabelecida antes do efetivo início da comunicação entre eles. Controla a divisão dos dados em pacotes, adiciona informações de controle de fluxo e sequenciamento, e coordena a confirmação e retransmissão de pacotes perdidos.

- **RTP – Real Time Protocol:** desenvolvido especialmente para transportar sinais cujo fator tempo é primordial, adiciona informações temporais e de sequenciamento a cada pacote, possibilitando seu reordenamento para reprodução de áudio e vídeo em tempo real.

Como bem colocado por Simpson e Greenfield (2007), considerando a hierarquia de rede estabelecida pelo modelo OSI os três protocolos anteriormente mencionados operam uma camada acima do protocolo IP, já que se baseiam em serviços de transporte de dados IP para efetivamente mover os dados de um dispositivo para outro na rede, como representado pela Figura 19:

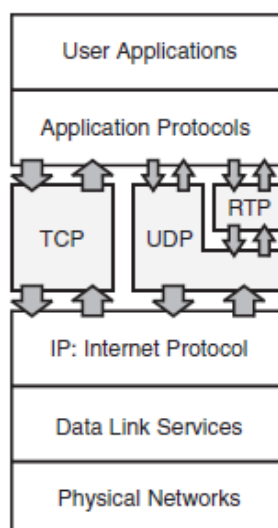


Figura 19: Hierarquia dos protocolos de transporte.

Fonte: IPTV and Internet Video, Simpson e Greenfield (2007).

A padronização EBU 3326 estabelece a utilização dos protocolos RTP e UDP para transporte de áudio, preferencialmente através de transmissões unidirecionais e com economia de informações nos cabeçalhos dos pacotes. A utilização do protocolo TCP é colocada como opcional pela padronização por não ser considerado adequado para transmissão de áudio, pois o mecanismo de retransmissão de pacotes perdidos tende a provocar atrasos e redução na eficiência da rede. Além disso, segundo a norma, o TCP possui menor prioridade nos roteadores que os protocolos RTP e UDP.

4.8.2 Codificação de Áudio

Os algoritmos de codificação de áudio (CODECS), também conhecidos por algoritmos de compressão, são utilizados com o objetivo de se reduzir a quantidade de bits necessária para representar um sinal de áudio. Um exemplo clássico de compressão de áudio é o formato MP3, capaz de compactar todo o conteúdo de um CD de música para um tamanho cerca de dez vezes menor, sem perda perceptível da qualidade sonora.

A tática dos CODECS consiste em explorar as limitações do ouvido humano para comprimir e codificar sinais de áudio sem que se perceba degradação na qualidade do som. Segundo Hoeg e Lauterbach (2003), isso é conseguido basicamente por meio de dois mecanismos: um que remove as redundâncias dos sinais de áudio através de correlações estatísticas e outro que considera fenômenos psicoacústicos como mascaramento temporal e espectral (Zwicker, 1967).

Simpson e Greenfield (2007) elencam as principais vantagens de se codificar sinais de áudio para distribuição em redes IP:

- Conteúdo comprimido pode ser transmitido através de redes menos velozes, que não teriam capacidade de trafegar os dados em seu formato original.
- A quantidade de canais de áudio transmitidos simultaneamente pela rede pode ser aumentada, em razão do ganho de escala proporcionado pela compressão dos sinais.
- Sinais de áudio comprimidos ocupam menos espaço nas mídias de armazenamento locais, tais como discos rígidos e servidores de backup, garantindo mais tempo de gravação.

Apesar das vantagens, Simpson e Greenfield (2007) lembram dos efeitos colaterais que os métodos de compressão podem causar, tais como a introdução de atrasos e ruídos nos sinais de áudio.

A escolha correta do método de compressão é de suma importância para o sucesso de um sistema de áudio sobre IP. A padronização EBU 3326 especifica quatro CODECS de áudio definidos pela norma como “obrigatórios”, por serem considerados de baixo custo e fácil implementação. São eles:

- **G.711:** utilizado na maioria dos sistemas de voz sobre IP a taxas de 64 kbit/s. O modelo u-law é o adotado nos Estados Unidos e Japão, enquanto o A-law se aplica à Europa e demais países. Utiliza como padrão períodos de 20 ms de áudio por pacote RTP, visando à melhor integração com os sistemas VoIP.
- **G.722:** muito comum e de fácil implementação, apesar da largura de banda de áudio limitada. Também utiliza taxa de bits de 64 kbit/s e 20 ms de áudio por pacote RTP.
- **ISO MPEG-1/2 Layer II:** formato de baixa complexidade e baixo custo de patente (direitos autorais). Proporciona boa qualidade de áudio a taxas de bits médias e altas, entre 32 e 384 kbit/s
- **PCM:** formato de áudio linear sem custos de patente (direitos autorais), baixa complexidade e isento de degradações decorrentes de processos de codificação em cascata. Quantização de 16, 20 e 24 bits e taxas de amostragem de 32 a 48 kHz.

Esses formatos devem estar presentes em todo e qualquer equipamento certificado de acordo com a especificação EBU N/ACIP. Além desses, são relacionados outros codecs de áudio apontados como “recomendados” (MPEG-4 AAC, MPEG-4 AAC-LD, MPEG-1/2 layer 3 a 32–320 kbps) e “opcionais” (MPEG-4 HE-AACv2, Enhanced APT-x, Dolby AC-3, AMR-WBp).

4.8.3 Tipos de Conexões IP

Segundo o documento EBU 3329, existem dois tipos possíveis de comunicação entre dispositivos em uma rede de áudio IP:

- **Unicast:** comunicação de um pra um. É o método de entrega de pacotes mais simples e mais utilizado em redes de dados, no entanto, pode se tornar inviável quando há necessidade de distribuir o mesmo conteúdo a diversos destinatários simultaneamente, já que o número de transmissões é igual ao número de clientes. Ou seja, cada transmissão é enviada exclusivamente para um único receptor, então, se vários clientes desejam receber o mesmo conteúdo, a fonte fica sobrecarregada ao ter que gerar unicasts dedicados a cada receptor. Neste caso, exige-se grande poder de processamento e largura de banda suficientemente alta para comportar o imenso tráfego de pacotes pela rede.
- **Multicast:** comunicação de um para muitos. É de fácil implementação, mas exige mecanismos de sinalização especiais e algoritmos de roteamento para controle de entrega. Além disso, requer suporte e configurações avançadas nos roteadores e switches da rede para replicação dos sinais transmitidos.

O método Multicast é o que melhor se aplica aos sistemas de áudio sobre IP, pois proporciona uma comunicação mais eficiente entre o transmissor e os dispositivos remotos, exigindo pouca largura de banda da rede. Neste caso, uma única fonte de áudio é disponibilizada para diversos clientes simultaneamente, assim como nos sistemas tradicionais de radiodifusão, onde um único sinal é transmitido a longas distâncias e qualquer pessoa com um aparelho de rádio pode captá-lo.

Simpson e Greenfield (2007) descrevem que em conexões Multicast um único sinal de áudio é enviado simultaneamente para múltiplos usuários, onde através de protocolos especiais, os roteadores e switches da rede são programados para fornecer aos receptores cópias fiéis do áudio transmitido pela fonte. Essas cópias

são feitas somente nos pontos necessários da rede, onde há solicitações de recebimento do conteúdo por parte dos clientes. Isso é possível por meio da atribuição de endereços especiais reservados aos pacotes multicast.

Segundo esses mesmos autores, sistemas multicast são unidirecionais e não possuem funcionalidades para recebimento de informações enviadas pelos clientes, assim, qualquer interação entre a fonte e os receptores deve ser gerenciada por outros tipos de mecanismos.

Por fim, cumpre ressaltar que redes públicas como a internet geralmente não estão habilitadas para conexões do tipo multicast, sendo sua aplicação restrita às redes locais (LAN). A Figura 20 ilustra a diferença no modo como os pacotes trafegam sob os métodos Unicast e Multicast.

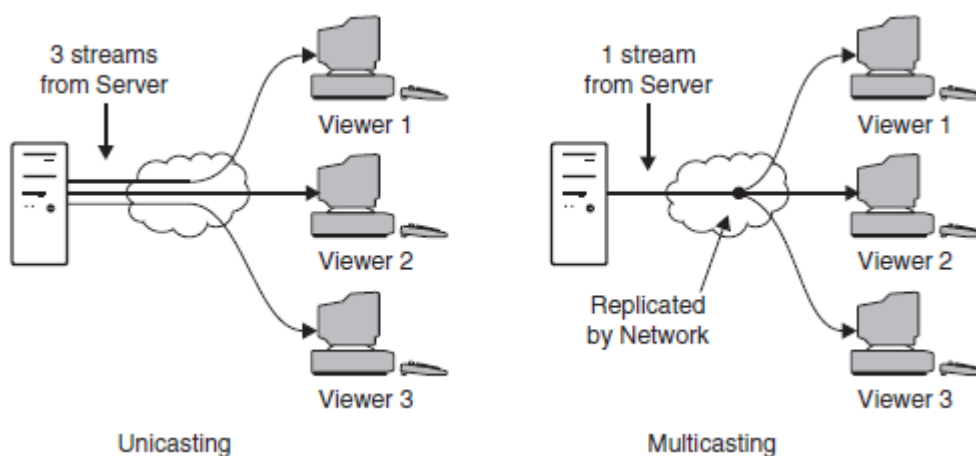


Figura 20: Unicast x Multicast.

Fonte: IPTV and Internet Video, Simpson e Greenfield (2007).

4.8.4 Protocolos de Sinalização

Protocolos de Sinalização tratam da configuração e negociação de parâmetros para realização de chamadas entre dispositivos. O protocolo SIP – *Session Initiation Protocol*, bastante utilizado em sistemas de telefonia VoIP, também foi o escolhido pelo EBU N/ACIP nas especificações de interoperabilidade dos equipamentos de áudio. Com isso, abre-se a possibilidade de integração entre telefones (fixos ou móveis) e sistemas de áudio sobre IP.

Conforme a padronização EBU 3326, o protocolo SIP utiliza a porta 5060, por meio da qual os terminais recebem as requisições de entrada, mantendo-se ativo somente durante a conexão, controle e finalização das chamadas. Quando combinado com o protocolo SDP – *Session Description Protocol*, o SIP agrega a funcionalidade de negociação do tipo de CODEC de áudio a ser utilizado na transmissão durante o estabelecimento da chamada.

De acordo com Church e Pizzi (2010), SIP deve ser utilizado como método de sinalização para links bidirecionais, enquanto em links unidirecionais, como é o caso das transmissões multicast, utiliza-se para anúncio das chamadas o protocolo SAPv1 – *Session Announcement Protocol*. Ainda segundo os autores, por padrão, as chamadas devem ser estabelecidas sob o mesmo CODEC de áudio em ambas as direções (transmissão e recepção), no entanto, é recomendado que o sistema suporte a utilização de CODECS distintos em cada sentido, bem como permita a mudança de CODECS durante conexões já estabelecidas.

4.9 QUALIDADE DO ÁUDIO

Em razão da natureza das redes IP, os pacotes de áudio que por elas trafegam estão sujeitos a problemas que podem ocasionar a degradação das informações sonoras, geralmente causadas por ruídos, atrasos de transmissão, erros de codificação e decodificação, defeitos no cabeamento, entre outros fatores. Problemas como esses podem gerar interferências que prejudicam a qualidade e fidelidade dos sinais de áudio transmitidos, facilmente percebido pela clareza e nitidez dos sons reproduzidos.

Segundo Simpson e Greenfield (2007), as imperfeições causadas nos sinais de áudio são decorrentes dos seguintes problemas de transmissão e codificação:

- **Perda de Pacotes:** é um dos erros mais comuns em sistemas de áudio IP. Pode ser causado por vários motivos, incluindo erros de bits que corrompem o cabeçalho dos pacotes, links sobrecarregados que forçam o descarte de pacotes nos roteadores, equipamentos de redes inadequados ou com defeito, entre outros. Os autores afirmam que este tipo de erro pode ser minimizado através do uso de práticas cuidadosas de projeto de sistemas e pelo rígido controle do tráfego de dados na rede.
- **Jitter:** ocorre quando os pacotes chegam ao seu destino de uma maneira não uniforme, num fluxo desordenado, adiantados ou atrasados em relação ao tempo previsto.
- **Erros de Bits:** ocorre quando os dados entregues ao destinatário são diferentes dos dados gerados na origem.
- **Ordenamento:** recepção desordenada dos pacotes de dados.
- **Fragmentação:** fragmentação de pacotes que excedem o tamanho máximo permitido.

Hoeg e Lauterbach (2003) afirmam que nos sistemas digitais de áudio sobre IP a qualidade final do áudio é determinada essencialmente pelos parâmetros de codificação e compressão (CODECS) da fonte transmissora, com pouquíssima

influência do meio de transmissão, ao contrário do que ocorria nos tradicionais sistemas analógicos. No entanto, isso significa também que, mesmo em um meio de transmissão perfeito, a qualidade do áudio pode ser limitada se os parâmetros operacionais dos CODECS não forem ajustados adequadamente.

Os autores entendem que nesta era de transmissões digitais a qualidade do áudio não é mais definida por critérios objetivos de medição, tais como relação sinal-ruído, linearidade, distorção ou resposta de frequências. A qualidade do áudio agora é definida em função do sinal original e as diferenças percebidas entre essa referência e o sinal codificado (comprimido) disponibilizado aos ouvintes.

As imperfeições introduzidas pelos CODECS nos sinais de áudio incluem distorção linear, ruído de quantização, ecos, limitações da faixa de áudio, alterações de timbre e da imagem estéreo. A Tabela 4 apresenta os defeitos e problemas mais recorrentes que contribuem para a degradação do sinal de áudio em sistemas de sonorização baseados em redes IP.

Tabela 4: Imperfeições causadas por técnicas de codificação e transmissão digitais.

Tipo	Descrição
Erro de Quantização	Defeitos associados com resolução insuficiente, por exemplo, distorção granular e alterações do nível de ruído.
Distorção de frequências	Ausência ou excesso de frequências graves, médias e/ou agudas.
Distorção de amplitude	Mudanças bruscas de amplitude e variações dinâmicas.
Efeito de modulação periódica	Variações periódicas de amplitude.
Efeito de modulação não-periódica	Deformação de períodos transientes.
Distorção Temporal	Pré-ecos e pós-ecos.
Ruído	Fantasmas e sons estranhos que não fazem parte do sinal original, como clicks, pops, glitches, etc.
Silêncio	Ausência de componentes sonoros do áudio original.
Crosstalk	Cruzamento de sons de sinais distintos.
Distorção da imagem espacial	Alterações de balanço, movimento, localização e espacialidade da imagem estéreo.

Fonte: Digital Audio Broadcasting, Hoeg e Lauterbach (2003).

5 CONCLUSÕES

Nem sempre uma nova tecnologia oferece considerável avanço técnico, facilidade de instalação e manutenção, maiores flexibilidade e escalabilidade e redução de custos quando comparada à tecnologia antecessora. Este é o caso dos sistemas de áudio sobre IP, que apresentam grande tendência de crescimento e aceitação no mercado de áudio profissional, tanto entre os profissionais da área quanto na indústria de conteúdo multimídia, ocasionando a migração e substituição dos antigos sistemas de sonorização analógicos. É uma tecnologia que já se encontra suficientemente consolidada e preparada para imediata implantação.

A grande convergência, popularização e barateamento das infraestruturas de redes de telecomunicações propiciam um ambiente totalmente favorável à implantação de sistemas de áudio sobre IP, assim como a crescente disponibilidade e ampliação da largura de banda que permite o tráfego de dados em altas velocidades para aplicações de tempo real, como é o caso da transmissão e distribuição de áudio.

Com redes estruturadas e bem gerenciadas, por meio de priorização de pacotes, balanceamento de carga, escolha correta de algoritmos de compressão e codificação, etc, é possível usufruir de todos os benefícios que esses sistemas proporcionam, contornando e controlando imperfeições dos sinais de áudio causados por problemas inerentes aos processos de digitalização e transmissão em redes de telecomunicações.

6 REFERÊNCIAS

AXIA AUDIO. **Introduction to Livewire: IP-Audio System Design Reference & Primer.** Cleveland OH USA.

AXIA AUDIO. **Multi-studio Acquisition and Installation Cost Comparison:** Purchase costs and installation expenses when comparing several facility solutions. Cleveland OH USA.

CLARK NOVAK. **IP-Audio Networks in the Real World:** How MPR, Univision And The University Of Indianapolis Are Using IP-Audio To Solve Problems. Cleveland OH USA.

DANIEL MINOLI. **IP multicast with applications to IPTV and mobile DVB-H.** John Wiley & Sons, 2008.

EBU Doc Tech 3326: **Audio contribution over IP:** Requirements for Interoperability. Geneva, April 2008.

EBU Doc Tech 3329: **A Tutorial on Audio Contribution over IP.** Geneva, May 2008.

HERVÉ BENOIT. **Digital television:** satellite, cable, terrestrial, iptv, mobile tv in the dvb framework. 3rd edition. Elsevier, Focal Press, 2008.

LAWRENCE HARTE. **Introduction to IP Audio;** Digitization, Compression and Transmission. Althos, May 17, 2007.

MIGUEL RATTON, **Dicionário de Áudio e Tecnologia Musical.** 1ª Ed., Editora Música & Tecnologia, 2004.

MIGUEL RATTON, **Fundamentos de Áudio.** 2ª Ed., Infomus Music Center Ltda., 2007.

SÓLON DO VALLE, **Manual Prático de Acústica.** 2ª Ed., Editora Música & Tecnologia, 2007.

STEVE CHURCH, SKIP PIZZI. **Audio over IP:** Building Pro AoIP Systems with Livewire. Elsevier, Focal Press, 2010.

TIELINE PTY. LTD. **The Audio over IP Instant Expert Guide.** Version 1.1, January, 2010.

WES SIMPSON, HOWARD GREENFIELD. **IPTV and Internet Video:** New Markets in Television Broadcasting. Elsevier, Focal Press, 2007.

WHEATSTONE CORPORATION. **What You Need to Know About Ethernet Audio.** Technical Document, 2009.

WOLFGANG HOEG, THOMAS LAUTERBACH. **Digital Audio Broadcasting:** Principles And Applications Of Digital Radio. Second Edition, John Wiley & Sons, 2003.